


# Incomplete Multi-view Clustering via Local Reasoning and Correlation Analysis

Xiaocui Li

Hunan University of Technology and  
Business, Xiangjiang Laboratory  
Changsha, Hunan, China  
xiaocuiworld@163.com

Guoliang Li

Hunan University of Technology and  
Business  
Changsha, Hunan, China  
liguoliang2111@163.com

Xinyu Zhang 

Hunan University of Technology and  
Business, Xiangjiang Laboratory  
Changsha, China  
zhangxinyu@whu.edu.cn

Yangtao Wang 

Guangzhou University  
Guangdong, Guangzhou, China  
ytaowang@gzhu.edu.cn

Qingyu Shi

Hunan University of Technology and  
Business, Xiangjiang Laboratory  
Changsha, Hunna, China  
qingyushi@hutb.edu.cn

Wei Liang

Hunan University of Technology and  
Business, Xiangjiang Laboratory  
Changsha, Hunna, China  
weiliang@csu.edu.cn

## Abstract

In recent years, incomplete multi-view clustering (IMVC) has attracted considerable attention for its ability to achieve effective clustering results through the integration of key information amidst missing view. However, the existing IMVC methods are still faced with 3 limitations: (1) They exhibit deficiencies in considering the weight distribution within views, (2) they ignore the varying contributions of different views to the common consistent representation, and (3) they struggle to sufficiently extract and recover the vital information within incomplete views. To address these limitations, we incorporate local reasoning and correlation analysis to design an incomplete multi-view clustering method (IMVCLRCA), which introduces a new strategy of feature learning and missing view recovery, fully exploiting local similarity and structural continuity within views and performing precise local reasoning recovery on missing data. By maximizing mutual information between views through contrastive learning, we achieve the consistent representation learning of multiple views. Furthermore, based on semantic consistency, we comprehensively consider the correlation between views, utilized a weight matrix to fuse cross-view data, and constructed a view with a correlation structure, ultimately obtaining a common consistent representation. We conduct extensive experiments on 4 public datasets including Caltech101-20, BBCSport, Scene-15, and LandUse-21. Experimental results demonstrate that IMVCLRCA has higher accuracy and robustness compared to the state-of-the-art IMVC methods. The anonymous code of this project is available on GitHub at <https://github.com/ggg2111/2025WSDM-IMVCLRCA>.



## CCS Concepts

• **Computing methodologies** → **Machine learning approaches; Feature selection; Computer vision representations.**

## Keywords

Incomplete multi-view clustering, Local information reasoning, Correlation analysis, Contrastive learning

## ACM Reference Format:

Xiaocui Li, Guoliang Li, Xinyu Zhang , Yangtao Wang , Qingyu Shi, and Wei Liang. 2025. Incomplete Multi-view Clustering via Local Reasoning and Correlation Analysis. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM '25)*, March 10–14, 2025, Hannover, Germany. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3701551.3703495>

## 1 Introduction

Multi-view data [1] describes the data from different sources and different feature spaces. For example, social data may include features like images and text, while bioinformatics data can cover multiple modalities such as facial, iris, and fingerprint recognition. Additionally, image data can be represented through various forms of description such as SIFT, LBP, and HOG. However, in the process of data collection, equipment failure and other reasons lead to missing multi-view data, which brings a series of problems and challenges to multi-view clustering. So incomplete multi-view clustering [2] emerges as the times require. IMVC can effectively integrate and utilize the relevant information between views to make up for the information vacancy caused by missing data, so as to achieve more accurate and robust clustering results.

In incomplete multi-view data [3], samples can be divided into complete samples, partial missing samples and complete missing samples according to the missing situation. The existing IMVC research mainly focuses on how to effectively fuse the limited and incomplete multi-view information, compensate or reconstruct missing data to improve clustering performance. Researchers have made various attempts, such as: Xue et al. proposed a clustering-induced adaptive structure enhancement network for incomplete multi-view clustering [4]; Yin et al. proposed incomplete multi-view clustering based on cosine similarity (IMCCS) [5]. Xie et al.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

WSDM '25, March 10–14, 2025, Hannover, Germany

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-1329-3/25/03

<https://doi.org/10.1145/3701551.3703495>

proposed an incomplete multi-view subspace clustering algorithm IMDF [6] based on adaptive instance-sample mapping and deep feature fusion. The above methods have improved the accuracy and robustness of incomplete multi-view clustering to a certain extent. However, they still have the following 3 shortcomings: (1) **They exhibit deficiencies in considering the weight distribution within views;** (2) **they ignore the varying contributions of different views to the common consistent representation;** (3) **they struggle to sufficiently extract and recover the vital information within incomplete views.**

To address the above problems, in this paper, we propose an Incomplete Multi-View Clustering via Local Reasoning and Correlation Analysis (IMVCLRCA), whose structural framework is illustrated in Figure 1. For Problem 1 and 2, we assign intra-view weights through a global correlation analysis algorithm, and fully fuse views to achieve inter-view correlation analysis and obtain a common consistent representation. For Problem 3, since adjacent data points often have similar features and there are often obvious spatial or sequential structures within samples, we propose a new view recovery strategy to fully exploit the local similarity and structural continuity to realize local information reasoning of views. At the same time, we maximize inter-view mutual information and information entropy through contrastive learning to achieve multi-view consistent representation learning.

The main contributions of this paper are summarized as follows:

- We propose a new local information reasoning strategy. The strategy directly exploits the inherent local similarity and structural continuity of views, and reconstructs the intra-view features through a self-supervised loss term.
- To fully consider the contribution of each view to the clustering, we propose a global correlation analysis algorithm, which performs inter-view correlation analysis while assigning intra-view weights, and fully integrates the information of each view to obtain a high-quality common consistent representation.
- We propose an IMVC method via local reasoning and correlation analysis. The method realizes missing data recovery and feature learning by mining potential features and connections within and between views, and achieves consistent representation learning through contrastive learning. We conduct experimental analyses on 4 datasets and the results demonstrate the superiority of IMVCLRCA.

## 2 Related Work

In this section, we briefly review two important topics related to our research, i.e., incomplete multi-view clustering and contrastive learning.

### 2.1 Incomplete Multi-view Clustering

The existing IMVC methods can be divided into: 1) IMVC based on matrix factorization[7 – 10], including non-negative matrix factorization (NMF) and singular value decomposition (SVD), etc., realizes low-rank approximation and feature extraction of data through matrix factorization technology. 2) IMVC based on kernel learning[11 – 14], including multi-view kernel alignment, multi-kernel learning, etc., maps multi-view data into high-dimensional

space by using kernel functions to capture nonlinear features, and integrates kernel matrices from different views to achieve incomplete multi-view clustering. 3) IMVC based on graph learning[15 – 19], including graph regularization, multi-view graph embedding, etc., captures the local and global relationships of data by constructing graph structures within and between views, and uses the properties of graph structures to deal with missing data, thus achieving incomplete multi-view clustering. 4) IMVC based on deep learning[20 – 24], including variational autoencoder (VAE), generative adversarial network (GAN), etc., processes complex and high-dimensional multi-view data through the powerful learning ability of deep neural networks, so as to improve the clustering performance.

Among the deep-learning-based IMVC studies, the most representative works include Incomplete Multi-View Clustering Based on Deep Semantic Mapping (IMVC-DSM) [25], Partial Multi-View Clustering Network Based on Consistent Generative Adversarial (PMVC-CGAN) [26], and Adversarial Incomplete Multi-View Clustering (AIMVC) [27]. Among them, IMVC-DSM solves the feature learning problem of multi-view data by exploring the partially aligned information in the available views, extracts deep features through DNN, and obtains a more reasonable structured representation through the regularization term of local graph, which is expressed as follows:

$$\min_{P_c, \{U^{(v)}, P_s^{(v)}\}_{v=1}^L} \sum_{v=1}^L \left\| \left[ A_c^{(v)}, A_s^{(v)} \right] - U^{(v)} \left[ P_c, P_s^{(v)} \right] \right\|_F^2 + \lambda_v \text{Tr} \left( P^{(v)} L_{Z^{(v)}} P^{(v)T} \right) \quad \text{s.t. } U^{(v)} \geq 0, P^{(v)} = \begin{bmatrix} P_c, P_s^{(v)} \end{bmatrix} \geq 0 \quad (1)$$

where  $A^{(v)} = \begin{bmatrix} A_c^{(v)}, A_s^{(v)} \end{bmatrix} = f \left( \vartheta^{(v)} X^{(v)} + b^{(v)} \right)$  represents the sample features generated by the  $v$ -th view in the DNN.  $L_{Z^{(v)}}$  is the pre-constructed nearest neighbor graph Laplacian matrix based on the available instances of the  $v$ -th view.  $\vartheta^{(v)}$  and  $b^{(v)}$  are network weights and biases, respectively.

PMVC-CGAN learns a shared low-dimensional representation and generates missing view data through GAN, and captures better common structure from part of the multi-view data to improve clustering performance. AIMVC infers missing data by integrating element reconstruction and GAN to capture overall structure and gain deeper semantic understanding, and designs pairwise clustering loss for better clustering structure.

### 2.2 Contrastive Learning

Contrastive learning learns the feature representation of data by constructing sample pairs and optimizing the closeness between similar samples and the spacing between dissimilar samples, so as to realize effective discrimination in the feature space, making similar samples gather while different samples disperse. In practice, contrastive learning extracts features through deep neural networks, optimizing models by loss function like InfoNCE to strengthen the ability to distinguish between positive and negative samples in the feature space. In MoCo [28], Kaiming He et al. regard contrastive learning as a dictionary query task, for example, for *query*  $q$  obtained from existing encoding and samples  $\{k_0, k_1, k_2, \dots\}$ ,  $\{k_0, k_1, k_2, \dots\}$  can be regarded as the *key* in the dictionary. If there

**Table 1: Description of the notation.**

Notation	Description
$X^{(v)}$	All samples under the $v$ -th view
$X_i^{(v)}$	The $v$ -th sample for the $v$ -th view
$Z^{(v)}$	All samples in the $v$ -th view after LIRM
$\tilde{X}^{(v)}$	All samples in the $v$ -th view after decoding
$\tilde{Z}^{(v)}$	All samples in the $v$ -th view after GCAA
$E^{(v)}(X^{(v)})$	The encoder for the $v$ -th view
$D^{(v)}(Z^{(v)})$	The decoder for the $v$ -th view
$LR^{(v)}$	The LIRM for the $v$ -th view
$CA^{(v)}$	Correlation analysis for the $v$ -th view
K	Class clusters

is only one *key* in the dictionary, i.e.  $k_+$  ( $k$  positive) matches *query*  $q$ , so  $q$  and  $k_+$  are considered positive sample pairs of each other, and the remaining *keys* are negative samples of  $q$ . Therefore, the loss function is defined as:

$$L_q = -\log \frac{\exp(q \cdot k_+ / \tau)}{\sum_{i=0}^k \exp(q \cdot k_i / \tau)} \quad (2)$$

To minimize the loss, *query*  $q$  is required to be similar to the unique positive sample  $k_+$  and dissimilar to all other negative samples key. Conversely, if the similarity between  $q$  and  $k_+$  is small or the similarity between  $q$  and other negative sample keys is large, the loss value will be relatively large, thus prompting the model to update parameters.

### 3 The Proposed Approach

#### 3.1 Notion

Given a multi-view dataset with  $N$  samples and  $V$  views  $X = \{X^{(1)}, X^{(2)}, \dots, X^{(V)}\}$ , where the multi-view dataset under the  $v$ -th view is denoted as  $X^{(v)} = \{x_1^{(v)}, x_2^{(v)}, x_3^{(v)}, \dots, x_N^{(v)}\}$ ,  $x_i^{(v)}$  represents the  $i$ -th sample feature under the  $v$ -th view. The key symbol definitions involved in this paper are shown in Table 1.

#### 3.2 Local reasoning and feature learning

In real application, the adjacent nodes or pixels of the multi-view data have similar textures or changes and exist global spatial or sequential structures, that is, local similarity and structural continuity. In this regard, we propose a new local information reasoning and feature learning strategy (LIRM), which processes data under each view, recovering missing data while fully learning features of views. Its core consists of six serialized convolution modules that perform convolution operations on input images by multiple filters, and each module is connected to a batch normalization layer to enhance stability and efficiency when processing high-dimensional data. Taking a two-dimensional image as an example,  $y_{ij}$  can be calculated as:

$$y_{ij} = \sum_{m=1}^M \sum_{n=1}^N x_{(i+m-1)(j+n-1)} \cdot k_{mn} \quad (3)$$

where  $x_{ij}$  represents the pixel value of the input image at position  $(i, j)$ ,  $k_{mn}$  represents the weight of the convolution kernel at position

**Table 2: The architecture of LIRM.**

Layer	Kernels	Size	Filling	Output
Input				$(B, 1, C, D)$
Conv1	$C$	$(C, 1)$	valid	$(B, C, 1, D)$
Permute				$(B, 1, C, D)$
BatchNorm				
Conv2	$C_q$	$(E_f * 0.1, 1)$	same	$(B, C_q, C, D)$
BatchNorm				
Conv3	$C_q$	$(E_f * 0.1, 1)$	same	$(B, C_q, C, D)$
BatchNorm				
Conv4	$C$	$(C, 1)$	same	$(B, C, C, D)$
BatchNorm				
Conv5	$1$	$(C, 1)$	same	$(B, 1, C, D)$
Conv6	$1$	$(C, 1)$	same	$(B, 1, C, D)$

$(m, n)$ , and  $y_{ij}$  represents the value of the output feature map at position  $(i, j)$ . If the convolution output of layer  $l$  is  $Y^{(l)}$ , then the input to layer  $l+1$  can be calculated as:

$$Y_{ij}^{(l+1)} = f \left( \sum_{m=1}^M \sum_{n=1}^N Y_{(i+m-1)(j+n-1)}^{(l)} \cdot k_{mn}^{(l)} \right) \quad (4)$$

where  $f$  represents the RELU activation function. Through the iteration of multi-layer convolution module, higher-level features (such as structure and shape) can be extracted from low-level local features (such as edges and corners) gradually, which is not only a hierarchical expression of local similarity, but also a mining method of spatial structural relationship. Therefore, for any partially missing sample, the local similarity and structural continuity can be fully utilized to effectively infer the missing data.

The module architecture is defined in Table 2. It should be noted that after *conv1* processing, the dimension of sample features will be transformed from  $(B, 1, C, D)$  to  $(B, C, 1, D)$ , which requires two-dimensional permutation on sample features to adapt to subsequent operations.

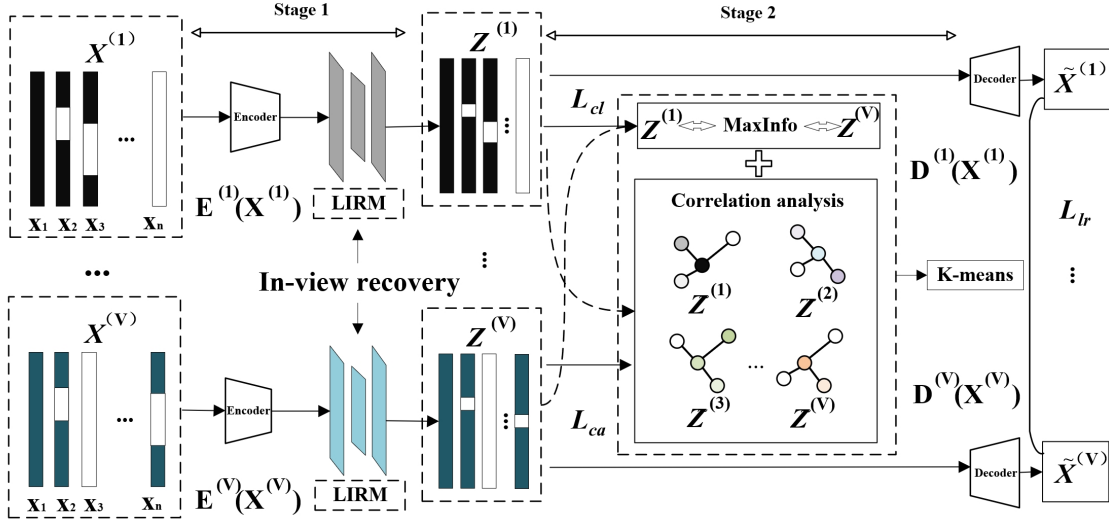
Since different datasets have different data characteristics, such as different view dimensions and sizes, the expansion factors  $E_f$  and  $C_q$  are set in the module to adjust the size and number of convolution kernels in different datasets, which makes the module flexible to adapt to the characteristics of different datasets and improves its ability to capture intra-view structural features.

Finally, the difference between the decoded views and the original views is minimized by the loss function  $\mathcal{L}_{lr}$ , which can be calculated as:

$$\mathcal{L}_{lr} = \sum_{i=1}^V \sum_{j=1}^N \left\| X_i^v - D^{(v)}(Z_i^v) \right\|_2^2 \quad (5)$$

$$Z_i^v = LR^{(v)} \left( E^{(v)}(X_i^v) \right) \quad (6)$$

From the above analysis, LIRM can improve the generalization ability of the network, effectively mine and recover useful information from incomplete multi-view data, and greatly improve the data utilization and recovery quality, which is of great significance for incomplete multi-view clustering in practical applications.



**Figure 1: The framework of the proposed IMVCLRCA. As shown, LRCAIMVC consists of three joint losses, i.e., local information reasoning, contrastive learning, and global correlation analysis.**

### 3.3 Global correlation analysis

To fully consider the contribution of each view to clustering and extract the key information in views, the sample features are weighted before the inter-view correlation analysis to distinguish the background and key information, and then the views are fused with each other through the weight matrix to obtain a common consistent representation.

Specifically, firstly, all sample features  $Z^{(v)}$  in a single view are input into L linear layers for weights learning, and query  $Q$  and keyword  $K$  are obtained through  $W_q$  and  $W_k$  respectively, where  $W_q, W_k \in R^{d \times d}$ ,  $Q, K \in R^{l \times d}$ ,  $l$  is the feature length,  $d$  is the feature depth. Subsequently, the query  $Q$  is normalized with the keyword  $K$  so that the norm of each feature vector is 1. The element-wise multiplication of the query matrix  $Q$  and  $W_e$  generates the original weights, which are then multiplied by a scaling factor to obtain the weights, thereby generating the global query weight vector  $e$ :

$$e = \frac{Q \cdot w_e}{\sqrt{d}} \quad (7)$$

where,  $w_e \in R^d$  is a learnable parameter vector. After the weight vector  $e$  is obtained, the normalization process is also performed to ensure that the weight distribution is reasonable, and then the weight  $e$  is used to perform weighted summation on the query  $Q$  to generate the global query vector  $q$ :

$$q = \sum_{i=1}^n e_i * Q_i \quad (8)$$

Where,  $e_i$  and  $Q_i$  are the parameters corresponding to the weight vector  $e$  and query  $Q$  under the  $i$ -th sample. After that, multiply the global query vector  $q$  with each vector of the keyword  $K$ , then after passing it through a linear layer  $W_g$ , add it and the original query  $Q$  together. Finally, input the output into the final linear layer  $W_o$

to obtain the weighted sample features:

$$Out = W_o \cdot (W_g \cdot (q \cdot K) + Q) \quad (9)$$

Finally, the difference between the fusion prediction results and the target sample features is minimized by the loss  $\mathcal{L}_{ca}$ .  $\mathcal{L}_{ca}$  can be calculated as:

$$\mathcal{L}_{ca} = \sum_{1 \leq v < i \leq V} \|\tilde{Z}^{(v)} - Z^{(i)}\|_2^2 \quad (10)$$

$$\tilde{Z}^{(v)} = CA^{(v)}(Z^i) \quad (11)$$

Through the above calculation, we complete the fusion of one view to another view and obtain  $\tilde{Z}^{(v)}$ . After the fusion of other views, we concatenate the fusion results with the original features to obtain the final common consistent representation of the current view. The same goes for the other views.

By weighted fusion of sample features, the algorithm can effectively process and fuse the effective information within and between views, thus improving the performance of IMNV.

### 3.4 Consistency represents learning

To maximize the consistency between the output features of the original view and the transformed view, contrastive learning is used to maximize the mutual information between views. Take two views as an example, sample features  $Z^{(1)}$  and  $Z^{(2)}$  under each view are obtained after the first stage of processing, then the mutual information between  $Z_i^{(1)}$  and  $Z_i^{(2)}$  is represented as  $I(Z_i^{(1)}; Z_i^{(2)})$ , and the information entropy of  $Z_i^{(1)}$  is represented as  $H(Z_i^{(1)})$ . In order to improve clustering performance, it is necessary to maximize the mutual information between samples from different views that is to maximize  $I(Z_i^{(v)}; Z_i^{(v)})$ , while according to information theory [29], the larger  $H(Z_i^{(v)})$  is, the more information the sample features contain. Therefore, the loss function  $\mathcal{L}_{cl}$  of contrastive

learning can be calculated as:

$$\mathcal{L}_{cl} = - \sum_{i=1}^N \left( I(Z_i^{(1)}; Z_i^{(2)}) + H(Z_i^{(1)}) + H(Z_i^{(2)}) \right) \quad (12)$$

Specifically, the input batch size  $B$  and feature dimension  $D$  are firstly obtained, and then the joint probability distribution is calculated to obtain the joint probability distribution matrix  $P \in R^{D \times D}$ .

$$P = \frac{1}{N} \sum_{i=1}^N Z_i^1 (Z_i^2)^\top \quad (13)$$

The symmetry and normalization are carried out by Eq. (14) to ensure that the joint probability distribution matrix  $P$  is symmetrical, and the sum of its elements is 1, so as to obtain the complete joint probability distribution matrix.

$$P = \frac{P'}{\sum_{i,j} P'_{ij}}, P' = \frac{P + P^T}{2} \quad (14)$$

Subsequently, the edge probability distribution  $P_i, P_j$  are calculated by (15), and the full loss  $\mathcal{L}_{ca}$  is obtained in the end.

$$P_i = \sum_j P_{ij}, P_j = \sum_i P_{ij} \quad (15)$$

$$\mathcal{L}_{cl} = - \sum_{i,j} P_{ij} (\log P_{ij} - \lambda \log P_i - \lambda \log P_j) \quad (16)$$

Through contrastive learning, the network can extract the essential features of data in the case of partial data missing to improve the quality of features from multi-view data, and promote the clustering effect.

Finally, we design a joint loss function, which comprehensively considers the local reasoning loss term, the correlation analysis loss term and the contrastive learning loss term. The joint loss function is defined as:

$$\mathcal{L}_{LRCA} = \mathcal{L}_{lr} + \lambda_1 \mathcal{L}_{ca} + \lambda_2 \mathcal{L}_{cl} \quad (17)$$

where parameter  $\lambda_1$  and  $\lambda_2$  are designed to balance  $\mathcal{L}_{lr}$  and  $\mathcal{L}_{ca}$ , which are all set to 0.1 in the experiment.

## 4 Experiments

### 4.1 Dataset

To comprehensively evaluate the effectiveness and applicability of the proposed method, four widely used public datasets are selected: Caltech101-20 [30], BBCSport [31], Scene-15 [32], and LandUse-21 [33], which are described in detail in Table 3. These datasets are from different application backgrounds, covering diverse image data from natural scenes to specific objects, and are suitable for testing the performance of multi-view clustering and information recovery algorithms when dealing with real complex data.

### 4.2 Comparison method

To comprehensively evaluate the methods proposed in this paper, 10 methods are selected and compared with our approach according to multiple dimensions such as release time and core technology. Meanwhile, according to their technical characteristics and processing strategies, they can be divided into the following categories:

- (1) Classical multi-view methods

**Table 3: Summarization of datasets.**

Dataset	Samples	Class	View	Feature
Caltech101-20	2386	20	6	48,40,254,1984,512,928
BBCSport	544	5	2	3183,3203
Scene-15	4485	15	3	20,59,40
LandUse-21	2100	21	3	20,59,40

DCCA [34] (ICML, 2013) enhances the traditional canonical correlation analysis through deep learning, enabling it to learn complex nonlinear mappings, thereby improving the correlation analysis efficiency of multi-view features. BMVC [35] (TPAMI, 2019) significantly reduces computational complexity and memory requirements through collaborative discrete representation learning and binary clustering structure learning, and utilizes an alternating optimization algorithm to ensure fast convergence.

- (2) Non-reasoning methods

AE2-Nets [36] (CVPR, 2019) enhance the ability of feature extraction and representation learning by nesting an autoencoder inside an autoencoder; PVC [37] solves the problem of partial view data missing by constructing a shared latent subspace. EERIMVC [38] (TPAMI, 2020) solves the complex problem of kernel matrix interpolation by associating each incomplete basis matrix generated by the incomplete view with the learned consensus matrix. PIC [39] uses spectral perturbation theory to complete the filling of similarity matrix.

- (3) Reasoning methods

UEAF [40] introduces error matrix and Laplacian regularization term to recover missing view; RecFormer [41] (TNNLS, 2023) uses a two-stage autoencoder network combined with self-attention structure to synchronously extract high-level semantic representations of multi-view data. DCP [42] (TPAMI, 2022) realizes information consistency and data recovery by jointly optimizing its contrastive learning loss and dual prediction loss. FCMVC-IV [43] (TIP, 2024) maintains a extensible consensus coefficient matrix, updates its knowledge with incomplete views passed in and designs a three-step iterative algorithm with linear complexity and provable convergence to fuse multi-view data.

### 4.3 Experimental results and analysis

In clustering evaluation, we selected accuracy (ACC), purity, normalized mutual information (NMI), V-measure (VM), adjusted Rand index (ARI) and Fowlkes-Mallows index (FMI) as evaluation indicators. Based on the experimental basis of [42], the IMVCLRCA proposed in this paper was tested and analyzed on various datasets with a missing rate of 50%. For all selected methods, we used the recommended network structure and parameters for fair comparison. The final experimental results are recorded in Tables 3-6, where the optimal values of each indicator are indicated in bold and the suboptimal ones are indicated by underscores. The experimental results show that:

- (1) The classical IMVC method is inferior to IMVCLRCA in all indicators. Firstly, both DCCA and BMVC can utilize the complementary information of multiple views to a certain extent, but both perform poorly in handling missing data, so their performance will

**Table 4: Experimental results on Caltech101-20.**

Method	ACC	Purity	NMI	VM	ARI	FMI
DCCA	38.60 ± 1.16	44.14 ± 3.54	52.54 ± 4.81	53.65 ± 4.99	29.79 ± 1.84	31.63 ± 2.81
BMVC	32.15 ± 2.92	49.07 ± 4.00	40.60 ± 4.33	41.23 ± 6.06	12.22 ± 0.89	22.95 ± 1.33
AE2-Nets	33.63 ± 3.03	48.22 ± 2.59	49.20 ± 4.24	54.71 ± 5.95	24.98 ± 1.85	31.49 ± 1.72
PVC	41.41 ± 2.95	39.90 ± 5.32	56.54 ± 2.04	60.42 ± 1.12	31.11 ± 0.81	34.28 ± 3.22
EERIMVC	40.66 ± 1.35	51.82 ± 1.63	51.38 ± 3.85	55.60 ± 4.01	27.91 ± 1.17	28.10 ± 3.39
UEAF	47.36 ± 2.50	46.26 ± 3.83	59.71 ± 4.03	57.93 ± 4.62	37.09 ± 2.19	31.84 ± 2.27
PIC	57.53 ± 4.06	60.21 ± 5.14	64.33 ± 5.31	65.80 ± 3.63	45.12 ± 1.88	49.98 ± 5.39
RecFormer	66.74 ± 6.95	69.17 ± 5.07	64.45 ± 5.48	63.65 ± 1.13	70.13 ± 2.59	67.42 ± 3.02
DCP	68.44 ± 1.14	71.08 ± 1.12	67.39 ± 2.02	<u>69.03 ± 7.09</u>	75.44 ± 2.46	<u>73.57 ± 2.70</u>
FCMVC-IV	<u>71.76 ± 3.39</u>	<u>75.48 ± 1.44</u>	<u>68.30 ± 8.11</u>	67.77 ± 3.66	<u>75.58 ± 1.02</u>	72.06 ± 0.68
IMVCLRCA	<b>75.89 ± 2.03</b>	<b>80.21 ± 2.41</b>	<b>70.42 ± 1.98</b>	<b>71.39 ± 1.79</b>	<b>79.15 ± 0.64</b>	<b>75.11 ± 1.02</b>

**Table 5: Experimental results on BBCSport.**

Method	ACC	Purity	NMI	VM	ARI	FMI
DCCA	27.43 ± 0.59	12.69 ± 1.00	21.78 ± 2.13	26.50 ± 2.13	23.38 ± 0.37	16.84 ± 0.40
BMVC	24.29 ± 0.84	22.06 ± 0.73	27.46 ± 1.95	24.62 ± 1.53	13.55 ± 1.97	12.01 ± 0.66
AE2-Nets	29.38 ± 1.12	27.74 ± 2.39	29.35 ± 0.24	30.19 ± 2.52	18.56 ± 2.41	12.89 ± 2.31
PVC	51.52 ± 1.42	45.26 ± 0.59	35.32 ± 2.29	40.48 ± 0.97	30.05 ± 1.53	26.12 ± 0.61
EERIMVC	35.43 ± 1.76	37.28 ± 3.00	34.19 ± 4.34	28.80 ± 0.31	23.90 ± 3.29	19.85 ± 5.07
UEAF	38.67 ± 0.41	30.04 ± 1.16	29.75 ± 0.94	26.45 ± 2.75	20.18 ± 1.18	16.48 ± 1.86
PIC	45.16 ± 0.31	35.20 ± 2.45	37.41 ± 4.29	33.07 ± 3.86	21.97 ± 2.25	20.62 ± 1.68
RecFormer	51.07 ± 1.37	44.28 ± 3.31	33.31 ± 5.53	36.73 ± 0.15	25.87 ± 0.37	22.99 ± 5.50
DCP	<u>60.37 ± 0.49</u>	<u>51.83 ± 1.69</u>	<b>41.20 ± 1.08</b>	39.26 ± 1.21	23.51 ± 2.04	20.44 ± 1.79
FCMVC-IV	58.42 ± 0.02	43.20 ± 1.07	37.56 ± 0.16	<u>41.13 ± 2.30</u>	<b>35.87 ± 1.65</b>	<u>27.29 ± 0.06</u>
IMVCLRCA	<b>62.81 ± 0.64</b>	<b>58.20 ± 0.36</b>	<u>40.25 ± 0.52</u>	<b>44.89 ± 2.10</b>	<u>35.74 ± 2.53</u>	<b>35.88 ± 0.22</b>

**Table 6: Experimental results on Scene-15.**

Method	ACC	Purity	NMI	VM	ARI	FMI
DCCA	31.83 ± 0.35	32.19 ± 0.60	33.18 ± 2.18	33.65 ± 1.02	14.95 ± 0.02	16.01 ± 1.76
BMVC	30.92 ± 1.50	31.90 ± 1.74	30.25 ± 0.87	29.32 ± 0.54	10.95 ± 1.84	12.47 ± 0.14
AE2-Nets	27.90 ± 0.33	29.14 ± 1.36	31.36 ± 0.06	33.19 ± 0.39	13.96 ± 0.08	14.89 ± 0.09
PVC	25.62 ± 1.31	24.62 ± 0.01	25.30 ± 1.61	28.46 ± 2.78	11.25 ± 0.08	13.21 ± 2.33
EERIMVC	33.11 ± 6.22	33.98 ± 4.13	32.12 ± 2.20	35.10 ± 4.41	15.90 ± 0.49	17.15 ± 2.41
UEAF	28.22 ± 0.34	28.01 ± 2.26	27.01 ± 0.01	25.56 ± 1.84	8.69 ± 0.96	11.48 ± 0.14
PIC	38.71 ± 3.73	36.74 ± 2.78	37.98 ± 0.25	41.77 ± 2.97	21.17 ± 3.21	23.62 ± 3.00
RecFormer	31.51 ± 5.37	33.28 ± 4.59	33.31 ± 4.11	30.83 ± 0.95	19.82 ± 1.24	22.41 ± 2.20
DCP	39.50 ± 0.71	42.73 ± 2.79	<u>42.35 ± 1.77</u>	<u>41.79 ± 2.17</u>	<u>23.51 ± 1.64</u>	25.87 ± 0.65
FCMVC-IV	<u>39.64 ± 2.57</u>	<u>43.20 ± 1.56</u>	38.45 ± 2.83	40.53 ± 1.25	22.66 ± 0.09	<u>26.02 ± 1.69</u>
IMVCLRCA	<b>42.70 ± 0.04</b>	<b>44.86 ± 0.78</b>	<b>44.93 ± 1.69</b>	<b>43.89 ± 0.21</b>	<b>25.24 ± 2.75</b>	<b>30.88 ± 0.43</b>

significantly decrease at higher missing rates. On the Caltech101-20 dataset, the ACC, Purity and NMI indexes of them were (38.60%, 44.14%, 52.54%) and (32.15%, 49.07%, 40.60%), respectively. Even DCCA’s ACC and Purity indicators in the LandUse-21 are only 14.09% and 9.82%. It can be seen that although the classical multi-view method can make use of multi-view data to a certain extent, it lacks the processing of missing data and its performance is very limited in the case of incomplete multi-view. Therefore, IMVCLRCA

makes up for the impact caused by missing views through local information reasoning, so as to achieve more accurate clustering results.

(2) Compared with the non-reasoning method, IMVCLRCA also shows excellent performance. The ACC and Purity indexes on Caltech101-20 dataset reach 75.89% and 80.21%, while AE2-Nets are only 33.63% and 48.22%, and for PVC, only 41.41% and 39.90%. By using a regularization method, EERIMVC strikes a balance between

**Table 7: Experimental results on LandUse-21.**

Method	ACC	Purity	NMI	VM	ARI	FMI
DCCA	14.09 ± 2.46	9.82 ± 1.54	20.06 ± 0.92	19.84 ± 0.58	3.40 ± 0.56	4.21 ± 0.76
BMVC	18.78 ± 0.16	14.11 ± 1.71	18.73 ± 2.58	17.90 ± 0.05	3.79 ± 0.63	4.39 ± 0.78
AE2-Nets	19.22 ± 0.53	16.02 ± 1.92	23.04 ± 0.87	22.13 ± 2.22	5.76 ± 0.95	6.48 ± 0.36
PVC	21.34 ± 1.24	17.45 ± 2.95	23.16 ± 1.33	24.57 ± 2.36	8.09 ± 1.37	7.75 ± 0.62
EERIMVC	22.15 ± 2.14	18.96 ± 3.37	25.20 ± 0.60	23.55 ± 1.93	9.09 ± 1.26	9.13 ± 0.09
UEAF	16.39 ± 1.74	14.07 ± 1.66	18.43 ± 0.49	19.29 ± 2.26	3.80 ± 0.56	5.60 ± 0.80
PIC	<b>23.60 ± 1.18</b>	20.98 ± 3.70	26.53 ± 2.96	24.81 ± 2.57	9.45 ± 1.66	8.72 ± 0.23
RecFormer	20.11 ± 3.44	18.06 ± 2.86	23.63 ± 2.50	23.98 ± 3.87	9.85 ± 1.11	10.24 ± 1.32
DCP	22.16 ± 0.53	<u>20.69 ± 1.85</u>	<u>27.00 ± 1.59</u>	<u>26.27 ± 0.59</u>	<u>10.39 ± 1.14</u>	11.03 ± 0.70
FCMVC-IV	20.93 ± 0.70	18.87 ± 1.31	24.12 ± 0.36	24.63 ± 1.24	10.21 ± 1.83	<u>11.79 ± 0.95</u>
IMVCLRCA	<u>22.63 ± 1.22</u>	<b>21.86 ± 0.36</b>	<b>28.71 ± 0.74</b>	<b>27.28 ± 1.96</b>	<b>10.42 ± 0.71</b>	<b>13.75 ± 0.26</b>

recovering missing data and realizing clustering, and achieves sub-optimal results in ACC and Purity (40.66%, 51.82%). However, on the LandUse-21 dataset, the ACC index of PIC is 0.97% (23.60%-22.63%) higher than that of IMVCLRCA. The reason is that PIC cleverly uses the spectral perturbation theory and performs well in dealing with the missing problem of similarity, but its comprehensive performance is still inferior to IMVCLRCA. It can be seen that the non-reasoning methods and IMVCLRCA have their own advantages and disadvantages in dealing with incomplete multi-view data, but the comprehensive performance of IMVCLRCA is still better than these methods on all datasets.

(3) As one of the reasoning methods, IMVCLRCA deals with incomplete multi-view data through data recovery and consistency enhancement, and shows the best performance on all datasets. Its values of ACC index outperforms the suboptimal method FCMVC-IV by 4.13%, 4.39%, 3.14% and 1.70%, respectively. It surpasses DCP by 7.45%, 2.44%, 3.20% and 0.47%, and surpasses UEAF by 28.53%, 24.14%, 14.48% and 6.24%. Although FCMVC-IV, DCP and UEAF can recover the missing views and enhance the consistency between views with different effective techniques, IMVCLRCA performs better when dealing with incomplete multi-view data because it realizes local reasoning and correlation analysis for missing views.

#### 4.4 Experimental results and analysis under different missing rates

In the previous subsection, IMVCLRCA achieves excellent performance with 50% missing rate. However, the missing degree of real data is often not fixed. In order to prove that IMVCLRCA can adapt to datasets with various missing degrees, this subsection evaluates IMVCLRCA under different missing rates on the Caltech101-20 dataset. Compared with DCP (22' TPAMI), EERIMVC (20' TPAMI) and BMVC (19' TPAMI), the results show that the robustness and accuracy of the proposed method are also better than other methods under different missing rates. The experimental results are shown in Figure 2.

According to the results, IMVCLRCA achieves the best clustering performance in all indicators under all missing rates. Especially in the case of high missing rate (50%-70%), the Purity and V-measure indicators are significantly higher than other methods, indicating

**Table 8: Ablation experiment results (unit: %).**

$\mathcal{L}_{lr}$	$\mathcal{L}_{ca}$	$\mathcal{L}_{cl}$	ACC	Purity	NMI	VM	ARI	FMI
✓			38.95	30.63	26.69	23.29	14.69	19.11
	✓		38.18	33.02	25.86	22.26	28.32	30.42
		✓	51.69	50.98	60.02	53.79	47.58	46.80
✓	✓		53.90	55.40	51.33	48.77	47.11	44.52
✓		✓	55.99	59.61	60.62	56.72	47.58	43.38
	✓	✓	61.23	64.30	65.84	62.81	59.03	57.48
✓	✓	✓	<b>75.89</b>	<b>80.21</b>	<b>70.42</b>	<b>71.39</b>	<b>79.15</b>	<b>75.11</b>

that the proposed method has higher accuracy in handling missing data.

At the same time, with the increase of the missing rate, although the clustering accuracy of each method decreases, the decline trend of IMVCLRCA is obviously gradual than that of other methods. In particular, it is significantly slower than BMVC under NMI, ARI and V-measure indicators, and significantly slower than EERIMVC under Purity indicators. It shows that our method can make full use of the consistency and difference information between view, and better adapt to real data with various missing degrees, and has superior robustness.

#### 4.5 Ablation experiments

In order to verify the role and importance of local information reasoning module ( $\mathcal{L}_{lr}$ ), global correlation analysis algorithm ( $\mathcal{L}_{ca}$ ) and contrastive learning loss term ( $\mathcal{L}_{lr}$ ) in our method, this subsection selected the Caltech101-20 dataset with 50% missing rate as the experimental object for ablation experiment, and the experimental results are shown in Table 8. From the results:

1) When the local information reasoning module and the global correlation analysis module are used alone, the overall performance of the method fails to achieve the expected effect. However, under the combination of the two, the performance is significantly improved: NMI reaches 65.84%, only 4.58% less than the best result; The ACC index reaches 61.23%, only 14.66% less than the best result. The experimental results show that the latter can effectively utilize the processing results of the former, and complement and enhance them on the basis of the former. This synergistic effect

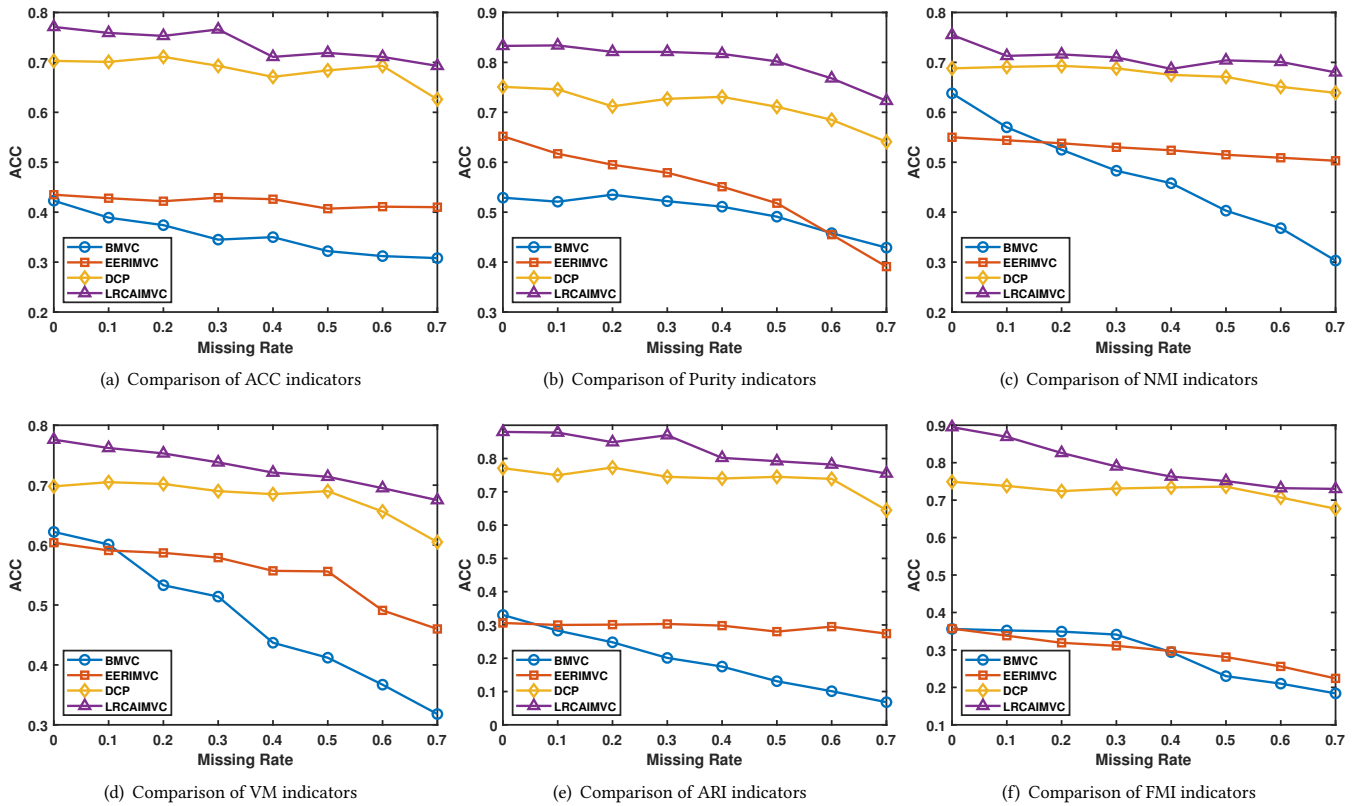


Figure 2: Comparison of indicators under different missing rate on Caltech101-20.

indicates that there is a structural and functional complementarity between two modules that allows them to significantly exceed their performance levels when acting jointly.

2) The performance of the inter-view contrastive learning module is very superior, and the ACC, NMI and VM indicators reach 51.69%, 60.02% and 53.79%, respectively, which proves the importance of consistent representation learning for IMVC. And it still has significant performance improvement when cooperated with other single modules, which proves that it can effectively use the processing results of view recovery and analysis to achieve a synergistic effect.

## 5 Conclusion and future work

This paper proposes an incomplete multi-view clustering method (IMVCLRCA) based on local information reasoning and global correlation construction, which improves the current IMVC research from the perspective of view recovery and correlation analysis inter-view. Different from other IMVC studies, IMVCLRCA fully considers the extraction and recovery of local features, and proposes a new strategy of feature learning and missing view recovery, which fully exploits the local similarity and structural continuity within views, and performs local information reasoning for missing data. In addition, IMVCLRCA maximizes the mutual information between views through contrastive learning to achieve consistent

representation learning of multi-view data. At the same time, IMVCLRCA also considers the different contributions of each view to the clustering, fully analyzes the correlation between various views based on semantic consistency, and fuses multi-view data by the weight matrix to obtain a common consistent representation to realize the correlation analysis between views. Finally, the experimental results on 4 datasets demonstrate the effectiveness of this method. In the future, we will conduct further research on different representation learning for multi-view data to achieve higher quality common consistent representation.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (No.62206092,62406082), the National Key Research and Development Program of China (No.2023YFC3306204), the Natural Science Foundation of Hunan Province (No.2023JJ40236, 2023JJ40239, 2022JJ40129, 2024JK2007), the MOE in China Project(23YJCZH183), the project of Xiangjiang Laboratory (No.XJ2022001, 22XJ03012, 22XJ03022), the Guangdong Basic and Applied Basic Research Foundation (No.2023A1515110650), the Guangzhou Science and Technology Planning Project (No.2024A03J0013), and the IoT Intelligent Sensing Support Project for Science and Technology Innovation Teams in Hunan Province.



## References

- [1] Zhou L, Du G, Lü K, et al. A Survey and an Empirical Evaluation of Multi-view Clustering Approaches[J]. *ACM Computing Surveys*, 2024, 56(7): 1-38.
- [2] Wen J, Zhang Z, Fei L, et al. A survey on incomplete multiview clustering[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 53(2): 1136-1149.
- [3] Tang J, Yi Q, Fu S, et al. Incomplete multi-view learning: Review, analysis, and prospects[J]. *Applied Soft Computing*, 2024: 111278.
- [4] Xue Z, Du J, Zheng C, et al. Clustering-Induced Adaptive Structure Enhancing Network for Incomplete Multi-View Data[C]//IJCAI. 2021: 3235-3241.
- [5] Yin J, Sun S. Incomplete multi-view clustering with cosine similarity[J]. *Pattern Recognition*, 2022, 123: 108371.
- [6] Xie M, Ye Z, Pan G, et al. Incomplete multi-view subspace clustering with adaptive instance-sample mapping and deep feature fusion[J]. *Applied Intelligence*, 2021, 51(8): 5584-5597.
- [7] Yin M, Liu X, Wang L, et al. Learning latent embedding via weighted projection matrix alignment for incomplete multi-view clustering[J]. *Information Sciences*, 2023, 634: 244-258.
- [8] Wen J, Xu G, Tang Z, et al. Graph regularized and feature aware matrix factorization for robust incomplete multi-view clustering[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.
- [9] Liu Z, Chen Z, Li Y, et al. IMC-NLT: Incomplete multi-view clustering by NMF and low-rank tensor[J]. *Expert Systems with Applications*, 2023, 221: 119742.
- [10] Lin J Q, Li X L, Chen M S, et al. Incomplete Data Meets Uncoupled Case: A Challenging Task of Multiview Clustering[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
- [11] Liu X. Incomplete multiple kernel alignment maximization for clustering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 46(3): 1412-1424.
- [12] Wang S, Cao J, Lei F, et al. Multiple kernel-based anchor graph coupled low-rank tensor learning for incomplete multi-view clustering[J]. *Applied Intelligence*, 2023, 53(4): 3687-3712.
- [13] Wu T, Feng S, Yuan J. Low-Rank Kernel Tensor Learning for Incomplete Multi-View Clustering[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2024, 38(14): 15952-15960.
- [14] Zhang G Y, Huang D, Wang C D. Unified and Tensorized Incomplete Multi-View Kernel Subspace Clustering[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2024.
- [15] Xue Z, Du J, Zhou H, et al. Robust diversified graph contrastive network for incomplete multi-view clustering[C]//Proceedings of the 30th ACM International Conference on Multimedia. 2022: 3936-3944.
- [16] Liu C, Wu Z, Wen J, et al. Localized sparse incomplete multi-view clustering[J]. *IEEE Transactions on Multimedia*, 2022, 25: 5539-5551.
- [17] Li X L, Chen M S, Wang C D, et al. Refining graph structure for incomplete multi-view clustering[J]. *IEEE transactions on neural networks and learning systems*, 2022, 35(2): 2300-2313.
- [18] Yang Z, Zhang H, Wei Y, et al. Geometric-inspired graph-based Incomplete Multi-view Clustering[J]. *Pattern Recognition*, 2024, 147: 110082.
- [19] Zhou W, Wang H, Yang Y. Consensus graph learning for incomplete multi-view clustering[C]//Advances in Knowledge Discovery and Data Mining: 23rd Pacific-Asia Conference, PAKDD 2019, Macau, China, April 14-17, 2019, Proceedings, Part I 23. Springer International Publishing, 2019: 529-540.
- [20] Liu C, Wu S, Li R, et al. Self-supervised graph completion for incomplete multi-view clustering[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(9): 9394-9406.
- [21] Xu J, Li C, Peng L, et al. Adaptive feature projection with distribution alignment for deep incomplete multi-view clustering[J]. *IEEE Transactions on Image Processing*, 2023, 32: 1354-1366.
- [22] Song P, Liu Z, Mu J, et al. Deep embedding based tensor incomplete multi-view clustering[J]. *Digital Signal Processing*, 2024, 151: 104534.
- [23] Xu G, Wen J, Liu C, et al. Deep Variational Incomplete Multi-View Clustering: Exploring Shared Clustering Structures[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2024, 38(14): 16147-16155.
- [24] Zhao Z, Cao Y, Xu Z, et al. Traffic emission estimation under incomplete information with spatiotemporal convolutional GAN[J]. *Neural Computing and Applications*, 2023, 35(21): 15821-15835.
- [25] Zhao L, Chen Z, Yang Y, et al. Incomplete multi-view clustering via deep semantic mapping[J]. *Neurocomputing*, 2018, 275: 1053-1062.
- [26] Wang Q, Ding Z, Tao Z, et al. Partial multi-view clustering via consistent GAN[C]//2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018: 1290-1295.
- [27] Xu C, Guan Z, Zhao W, et al. Adversarial incomplete multi-view clustering[C]//IJCAI. 2019, 7: 3933-3939.
- [28] He K, Fan H, Wu Y, et al. Momentum contrast for unsupervised visual representation learning[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 9729-9738.
- [29] Thomas M, Joy A T. Elements of information theory[M]. Wiley-Interscience, 2006.
- [30] Li Y, Nie F, Huang H, et al. Large-scale multi-view spectral clustering via bipartite graph[C]//Proceedings of the AAAI conference on artificial intelligence. 2015, 29(1): 2750-2756.
- [31] Greene D, Cunningham P. Practical solutions to the problem of diagonal dominance in kernel document clustering[C]//Proceedings of the 23rd international conference on Machine learning. 2006: 377-384.
- [32] Fei-Fei L, Perona P. A bayesian hierarchical model for learning natural scene categories[C]//2005 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2005, 2: 524-531.
- [33] Yang Y, Newsam S. Bag-of-visual-words and spatial extensions for land-use classification[C]//Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. 2010: 270-279.
- [34] Andrew G, Arora R, Bilmes J, et al. Deep canonical correlation analysis[C]//International conference on machine learning. PMLR, 2013: 1247-1255.
- [35] Zhang Z, Liu L, Shen F, et al. Binary multi-view clustering[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2018, 41(7): 1774-1782.
- [36] Zhang C, Liu Y, Fu H. Ae2-nets: Autoencoder in autoencoder networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 2577-2585.
- [37] Li S Y, Jiang Y, Zhou Z H. Partial multi-view clustering[C]//Proceedings of the AAAI conference on artificial intelligence. 2014, 28(1): 1968-1974.
- [38] Liu X, Li M, Tang C, et al. Efficient and effective regularized incomplete multi-view clustering[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2020, 43(8): 2634-2646.
- [39] Wang H, Zong L, Liu B, et al. Spectral perturbation meets incomplete multi-view data[C]//Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. 2019: 3677-3683.
- [40] Wen J, Zhang Z, Xu Y, et al. Unified embedding alignment with missing views inferring for incomplete multi-view clustering[C]//Proceedings of the AAAI conference on artificial intelligence. 2019, 33(01): 5393-5400.
- [41] Liu C, Wen J, Wu Z, et al. Information recovery-driven deep incomplete multiview clustering network[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [42] Lin Y, Gou Y, Liu X, et al. Dual contrastive prediction for incomplete multi-view representation learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(4): 4447-4461.
- [43] Shu Z, Li B, Mao C, et al. Structure-guided feature and cluster contrastive learning for multi-view clustering[J]. *Neurocomputing*, 2024, 582: 127555.